Express Mail N . EV231274929US
Date of Deposit: December 30, 2003

PATENT APPLICATION
Attorney Docket No: 19904-006 (BREESE-6) CON

OPTIMAL VIDEO DECODER BASED ON MPEG-TYPE STANDARDS

CLAIM OF PRIORITY

This application claims priority under 35 U.S.C. § 119(a) to French patent application 99 07443, filed June 11, 1999.

TECHNICAL FIELD OF THE INVENTION

The invention concerns the display of animated images, in particular decompression of digital data which incorporate these images using optimized methods.

BACKGROUND OF THE INVENTION

With the appearance of the most recent digital technologies, and the ever-increasing need for speed and storage space, compression cannot be avoided for mainstream applications. Examples include digital cameras which code images in JPEG, digital camcorders which compress DV format sequences, an M-JPEG derivative, or digital television and DVD, which have adopted the MPEG-2 compression format, in addition, of course, to the Internet, in which images and sequences are sent in compressed form.

In certain cases, the user requires very high quality (photo, camcorder), implying very low rates of compression. In other cases, excessive transfer time prevents acceptable quality. It is, therefore, necessary to improve sequence decoding to permit either better quality at an equivalent rate, or a weaker rate with equal or superior quality.

Different animated image compression standards have been proposed, but only the MPEG standard has really taken hold. This standard for the compression and decompression of animated images leads to the appearance of block effects.

For example, European patent EP539833 concerns a process designed to produce a compressed video data representation which can be displayed on a video screen after decompression according to a number of hierarchical scales of image and/or quality resolution, including phases which consist of:

- providing video image element data signals indicating block units in space or macro-blocks, which associate the information concerning the compressed image data with a group of coding attributes, including coding decisions, movement compensation vectors, and quantification parameters, and
- producing for each of these macro-blocks a macro block which is placed on the corresponding scale for each scale of this multiplicity so that the same coding attributes are shared by these scaled macro-blocks.

Methods enabling decompression errors to be corrected have been proposed by previous researchers. These methods primarily concern techniques used after the process of image decompression itself and slow down this compression. These methods do not take the quantifier into account and do not permit the binary frame to be retained after recompression, which has the effect of degrading image quality with each compression. A goal of the invention is to propose a process for improving image quality during decompression.

SUMMARY OF THE INVENTION

The invention concerns a process for decompression of compressed animated images with a method including treatment of images in blocks and containing a digital data recomposition phase defining predefined forms, a phase modeling the movement of these forms using a process of prediction, interpolation and temporal compensation, an image composition phase from reconstructed elements of JPEG or MPEG type motion. The form recomposition phase includes a process for separating fixed forms from mobile forms, a process for recording digital data corresponding to the fixed forms treated by a filter which is not separable from the processes implemented in the recomposition phase

in a first specific memory unit and digital data corresponding to mobile forms in a second specific memory unit.

Beneficially, the digital filter is irreducible and does not contain dissociable filters. In one variant, the filter eliminates the block effect on the background image. It can regularize the background image. Beneficially, the quantification interval used during compression of the background image is stored and projected on the quantification interval.

In one variant, reconstruction of the elements uses quantification parameters previously defined by the coder during image compression. These parameters are linked to the image photography methods and permit decompression to be adapted based on these methods. This permits taking account of the compression characteristics and improving image decompression. In one variant, the quantification parameters are defined by the transfer function of the methods of acquisition and storage of animated images.

Beneficially, a second digital filter separates and identifies the mobile elements into mobile objects moving in a sequence, in accordance with the evolution of predetermined digital criteria, such as the geometry of mobile objects, movement of mobile objects, or spatial segmentation of mobile objects. Temporal averaging can also take place with compensation for the movement of each identified object. In one variant, the filter eliminates the block effect from objects. According to this process, the objects identified can also be regularized.

Beneficially, the quantification interval serving to compress the animated sequence is stored and is projected on the quantification interval. During display, the mobile objects and averaged representation are superimposed in the fixed image time.

Preferably, the parameters specific to each object identified are stored separately in order to treat each object differently.

The invention also concerns a device for decompression of compressed animated images with a method including image treatment in blocks and a digital data recomposition phase defining predefined forms, a movement modeling stage of these

forms using methods for prediction, interpolation and temporal compensation, an image composition phase from reconstructed elements of JPEG or MPEG type motion. It includes methods for separating fixed forms from mobile forms, and methods for recording digital data corresponding to the fixed forms treated by a filter which is not separable from the methods implemented in the recomposition phase in a first specific memory unit, and digital data corresponding to mobile forms in a second specific memory unit.

This device also beneficially includes irreducible digital filtering methods, which cannot be decomposed into a sequence of filters independent of one another.

It includes, preferably, storage methods for the types of images compressed.

One variant of this device includes a detachable medium and can be made with an independent chip or a graphics memory card that can be inserted into a computer. This device can also be inserted without being separated into a computer or into any type of electronic apparatus permitting image display.

This device can also be comprised using an independent software module from the software present in a calculator memory.

The invention consists of a decoding process for reducing both the block effects and defects linked to degradation of the sequence media.

This method treats the problem spatially and temporally, obtaining significant improvement in sequence quality, and is based on two ideas:

- simultaneously treating the problems of block suppression and movement segmentation,
- integrating the notion of object, in order to permit a different approach for treatment of the background and of each object.

In addition, this method presents the particular characteristic of effectively treating the problem of "drop-out", independent of the original sequence format: the image blocks lost to acquisition by the camera, or during transmission are perfectly restored; and defects such as abrasions and threads are removed from digital film. In

addition, it is possible to integrate the process of accounting for the objective transfer function of the camera, or the projector, into the decoding process to obtain a more precise restitution. Finally, the scheme proposed remains valid within the framework of MPEG-4 decoding.

The method proposed uses an object approach with two distinct phases. Firstly, the sequence background is isolated and the block effects in it are suppressed. The objects are slowly isolated from the background, benefitting from a more precise representation at each stage. Each object is then treated independently, according to its own characteristics, and then finally projected on the estimated background image to reconstruct the sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a flow chart of the method of the invention.

DETAILED DESCRIPTION OF THE INVENTION

FIG. 1 represents the stages of this process.

A first stage (1) consists of estimation and background image treatment, in addition to identification of the mobile elements. A map of these mobile elements is transmitted by the treatment of these elements.

Stage (2) consists of pretreating this map by labeling and completing each element.

Stage (3) consists of spatial segmentation of the different elements permitting identification of the different mobile objects. This stage also permits estimation of this movement, and follow-up of objects during the sequence.

Stage (4) specifically treats each object identified according to the methods explained below.

Stage (5) consists of the reconstruction of the sequence which permits obtaining the decoded sequence.

Estimation of the background is considered to be an inverse problem. Let p_k be N images of the MPEG or M-JPEG sequence containing the block effects. The background is simultaneously estimated, and the sequence of mobile objects, termed c_k.

We want c_k=0 if the point belongs to a mobile object, otherwise c_k=1. We look for

$$f^* = \arg\min(J_\lambda)$$

with as criterion

$$J_\lambda(f, c_1, \ldots, c_N) = J_1(f, c_1, \ldots, c_N) + \lambda^2 J_2(f) + \gamma^2 J_3(c_k)$$

with

$$J_1(f, c_1, \ldots, c_N) = \sum_{k=1}^{N} \int_\Omega c_k^2\, (f - p_k)^2 + \alpha_c \sum_{k=1}^{N} \int_\Omega (c_k - 1)^2$$

which causes spatiotemporal segmentation using N consecutive images of the sequence, and

$$J_2(f) = \int_\Omega \phi_1(\|\nabla f\|) \qquad (1)$$

$$J_3(c_k) = \int_\Omega \phi_2(\|\nabla c_k\|) \qquad (2)$$

the regularization terms which a priori contain the solution. ϕ_1 and ϕ_2 are potential functions which maintain the discontinuities in the image. Parameter α_c determines the importance granted to the background: the smaller α_c is, the more mobile objects are detected.

Relative to $J_1(f)$, if p_k is far away from the current estimate f, c_k must be small: the object is moving.

This comprises a traditional approach for spatiotemporal segmentation of sequences. However, this method does not affect the block effects resulting from the DCT, and does not take into consideration coder characteristics. The treatment specific to the invention solves this problem.

To take account of the quantifier and simultaneously suppress the block effects during extraction of f background, the new criterion is minimized:

$$J(f, c_1, \ldots, c_N) = J_1(f, c_1, \ldots, c_N) + \lambda_1^2 J_2(f) + \gamma_1^2 J_3(f) + \eta_1^2 J_3(f) + \mu_1^2 J_3(f) \qquad (3)$$

with

$$J_4(f) = \int_\Omega \Psi \left(\frac{|Rf|}{\delta} \right) \qquad (4)$$

where R is the transformation into wavelets, Ψ a potential function, and δ a threshold dependent upon block effect amplitude. The value of c_l specifies which wavelet coefficients are to be thresholded. A soft thresholding in the spatio-frequential wavelet area is then performed.

Specific knowledge of the quantification matrix used during coding permits each pixel from the reconstructed sequence to be restricted to an interval corresponding to the quantification interval.

Quantification is a discretization operation which transforms a continuous group of sample values to a discrete group. It can be performed on a single sample at the same time (scalar quantification) or several samples assembled in blocks (vector quantification). The restriction corresponding to this projection is:

$$J_5(f) = \tfrac{1}{4} \int_\Omega \left(\left| Df - p_k + \frac{q}{2} \right| - Df + p_k - \frac{q}{2} \right)^2 + \tfrac{1}{4} \int_\Omega \left(\left| Df - p_k - \frac{q}{2} \right| - Df + p_k - \frac{q}{2} \right)^2$$

with D being the DCT operator, p_k the quantified DCT coefficient, q the quantification step for the pixel considered.

The f^* optimal solution for this minimization problem is found for $\frac{\partial J_{\lambda 1}}{\partial f} = 0$, equivalent to:

$$\sum_{k=1}^{N} c_k^2(f - p_k) - \lambda_1^2\, div\left(\frac{\phi_1'(|\nabla f|)}{|\nabla f|}\nabla f\right) + \eta_1^2 R^T \frac{\Psi'\left(|Rf|/\delta\right)}{\left(|Rf|/\delta\right)} Rf + \mu_1^2 D^T \kappa(f) = 0 \qquad (5)$$

with

$$\kappa(f) = \begin{cases} Df - p_k + \frac{q}{2} & \text{if} \quad Df < p_k - \frac{q}{2} \\ Df - p_k - \frac{q}{2} & \text{if} \quad Df > p_k + \frac{q}{2} \\ 0 & \text{if} \quad Df \in \left[p_k - \frac{q}{2}; p_k + \frac{q}{2}\right] \end{cases}$$

Optimal c_k objective cards are obtained for $\frac{\partial J_{\lambda 1}}{\partial c_k} = 0$, yielding the following equation:

$$\sum_{k=1}^{N} c_k (f - p_k)^2 + \alpha_c \sum_{k=1}^{N} (c_k - 1) - \gamma_1^2 div\left(\frac{\phi_1'(|\nabla c_k|)}{|\nabla c_k|}\nabla c_k\right) = 0 \qquad (6)$$

The problem is solved with two successive optimizations:

Minimization of (5) in f, for c_k given, → f^*

Minimization of (6) in c_k, for f^* given → c_k^*

These two optimizations are then iterated by searching for a new f^* background, followed by new c_k, until the convergence of the solution.

The criterion is resolved using a semi-quadratic resolution algorithm described in the article *Deterministic Edge-Preserving Regularization in Computed Imaging,* 5(12) IEEE Transaction on Image Processing (February 1997), based on alternating minimizations. Other methods may also be used.

This new criterion therefore suppresses the background block effects, and simultaneously segments moving objects.

This criterion provides a sequence of moving card elements. In order to be able to treat each element separately, they must be spatially isolated from one another. However, the more numerous the block effects are in the original sequence, the more the c_k cards present false or poor quality information. For example, a DCT block whose intensity changes from one image to the next may be thought of as a moving object. Several pre-treatments are therefore necessary before isolating each element:

- Thresholding of the c_k card. The values with intensity less than a given threshold are brought to 0, and the others to 1.
- Mathematical closure and filling of each object. Mathematical closure occurs, in other words dilatation followed by erosion, by a structuring element of size nxn, preferably with n=3. The element is filled using a traditional image path method. Other methods can also be used, such as active geodesic contours.
- Mathematical opening and suppression of certain objects. The opening consists of making an erosion followed by a dilatation, to suppress the false elements coming from the DCT blocks. Other methods can also be used.

From these c_k it is possible to label each element, isolate them from one another and consider them as objects. Henceforth, each treatment described will be completed independently on each object.

For each object in the sequence, certain characteristics will be determined which will permit a detailed and adapted treatment:

- Evolution of the shape, average height and size, position, barycenter, in the sequence.
- Object spatial segmentation information. A given object is spatially segmented to determine the different zones it contains (discontinuities, homogeneous zones...). Traditional methods for spatial segmentation of fixed images can be used.

- Estimation of object movement using traditional "block-matching" methods, or optical flow. This estimation of movement provides a movement vector $d=(dx_i, dy_i)$ for each object, and for each image i of the sequence.

Once each object has been isolated, and its movement determined, its treatment can be customized to suppress the block effects it contains. This phase may be performed in parallel on each object, to optimize speed of execution.

For each object, we look for:

$O_k^* = \arg\min(J_{\lambda 2})$

with

$$J_\lambda(O_k) = J_1(O_k) + \lambda_2^2 J_2(O_k) + \eta_2^2 J_3(O_k) + \mu_1^2 J_4(O_k) \qquad (7)$$

where

- $J_1(O_k) = \sum_{i=n}^{n} \int_\Omega (c_k - 1)^2 (O_k - p_{k+i}(x + dx_{k+i}, y + dy_{k+i}))^2$ is a temporal

averaging of the object, with compensation for movement. The value of n depends upon the object characteristics, in particular its non-stationary nature. The more rapidly the object evolves over time, the smaller the n chosen will be.

- $J_2(O_k) = \int_\Omega (c_k - 1)^2 \phi_3(\|\nabla O_k\|)$ regularizes the object. λ_2 is adaptive; it

depends upon the spatial segmentation chosen to determine the different object zones, and permits customizing the object treatment.

- $J_3(O_k) = \int_\Omega (c_k - 1)^2 \Psi\left(\frac{|RO_k|}{\delta}\right)$ suppresses the block effects on the object.

- $$J_4(O_k) = \tfrac{1}{4}\int_\Omega (c_k - 1)^2 \left(\left|DO_k - p_k + \frac{q}{2}\right| - DO_k + p_k - \frac{q}{2}\right)^2$$

$$+ \tfrac{1}{4}\int_\Omega (c_k - 1)^2 \left(\left|DO_k - p_k - \frac{q}{2}\right| - DO_k + p_k - \frac{q}{2}\right)^2$$

permits restricting each pixel of each object to the quantification interval, to reduce quantification noise on the object.

An optimal solution O_k^* is obtained for $\frac{\partial J_{\lambda_2}}{\partial O_k} = 0$, equivalent to:

$$(c_k - 1)^2 \left(\sum_{i=-n}^{n} \left(O_k - p_{k+i}\right) - \lambda_2^2\, div\left(\frac{\phi_3'(|\nabla O_k|)}{|\nabla O_k|} \nabla O_k \right) + \eta_2^2 R^T \frac{\Psi'\left(\frac{|RO_k|}{\delta}\right)}{\frac{|Rf|}{\delta}} RO_k + \mu_2^2 D^T \kappa(O_k) \right) = 0 \qquad (8)$$

with

$$\kappa(f) = \begin{cases} DO_k - p_k + \frac{q}{2} & \text{if} \quad DO_k < p_k - \frac{q}{2} \\ DO_k - p_k - \frac{q}{2} & \text{if} \quad DO_k > p_k + \frac{q}{2} \\ 0 & \text{if} \quad DO_k \in \left[p_k - \frac{q}{2} ; p_k + \frac{q}{2}\right] \end{cases}$$

The method of resolution used to solve the equation (8) is identical to that used above.

The method presented may be simplified, in order to reduce its complexity, and therefore calculation time.

Simplification during estimation of background and c_k.

A first simplification consists of setting η_1=0 in the equation (3). The wavelet coefficient thresholding can in this case be used as a pretreatment for each image entering the sequence. μ_1=0 is also posited in (3), and the interval restriction can be implemented by projection on the quantification intervals. The result of this simplification is a significant decrease in calculation time, at the cost of a slight decrease in quality.

The second simplification consists of suppressing regularization on the c_k, or positing γ_1=0 in (3). To obtain the c_k sequence, $\frac{\partial J_{\lambda_1}}{\partial c_k}\Big|_{\gamma_1 = 0} = 0$ is solved (*cf.* equation (6)).

An explicit formula $c_k^* = \frac{\alpha_c}{\alpha_c + (f - p_k)^2}$ is then obtained which permits calculation of the sequence of moving objects.

Equation (7) can be simplified:

- by positing λ_2^2 in (7), object regularization is suppressed. In this case, only temporal averaging of the object occurs, with compensation for movement.
- by positing $\eta_2 = 0$ in the equation (7) and performing thresholding as a pretreatment of each object.
- by positing $\mu_2 = 0$ in (7), and performing projection on the qualification intervals of each object.

By totaling some of these simplifications, the algorithm becomes quick, and may be adapted to real time applications.

The decoded sequence $\tilde{p}$ is reconstituted by using a background image over a duration of N images, and projecting into it the M objects:

$$\tilde{p}_k = c_k^{*2} f^* + (c_k^* - 1)^2 O_k^* \tag{9}$$

If, for a given pixel, you are on an object, $c_k^*=0$, the O_k pixel is then projected, otherwise $c_k^*=1$ and the pixel is projected from background f^*.

The details of one or more embodiments of the invention are set forth in the accompanying description above. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described. Other features, objects, and advantages of the invention will be apparent from the description and from the claims. In the specification and the appended claims, the singular forms include plural referents unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All patents and publications cited in this specification are incorporated by reference.

The foregoing description has been presented only for the purposes of illustration and is not intended to limit the invention to the precise form disclosed, but by the claims appended hereto.